

AD-779 290

**EVACUATION OF A YULE PROCESS WITH
IMMIGRATION**

Mark Brown, et al

George Washington University

Prepared for:

Office of Naval Research

29 April 1974

DISTRIBUTED BY:



**National Technical Information Service
U. S. DEPARTMENT OF COMMERCE
5285 Port Royal Road, Springfield Va. 22151**

UNCLASSIFIED

Security Classification

AD-779290

DOCUMENT CONTROL DATA - R & D

Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified

1. ORIGINATING ACTIVITY (Corporate author) THE GEORGE WASHINGTON UNIVERSITY DEPARTMENT OF STATISTICS WASHINGTON, D.C. 20006		2a. REPORT SECURITY CLASSIFICATION
		2b. GROUP
3. REPORT TITLE EVACUATION OF A YULE PROCESS WITH IMMIGRATION		
4. DESCRIPTIVE NOTES (Type of report and inclusive dates) TECHNICAL REPORT		
5. AUTHOR(S) (First name, middle initial, last name) MARK BROWN, SHELDON ROSS, RICHARD SHORROCK		
6. REPORT DATE April 29, 1974	7a. TOTAL NO. OF PAGES 18	7b. NO. OF REFS 3
8a. CONTRACT OR GRANT NO. N00014-67-A-0214-0015	9a. ORIGINATOR'S REPORT NUMBER(S) 19	
b. PROJECT NO. NR-042-267	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)	
c.		
d.		
10. DISTRIBUTION STATEMENT Unlimited. Reproduction in whole or in part is permitted for any purpose of the United States Government.		
11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY OFFICE OF NAVAL RESEARCH STATISTICS & PROBABILITY PROGRAM ARLINGTON, VIRGINIA 22217	
13. ABSTRACT		
<p>A correspondence is obtained between the epochs of a Yule process and orders statistics from an exponential distribution. This correspondence is used to obtain some properties of the Yule process, which are in turn used to solve an optimization problem. This problem involves the finding of an optimal evacuation time under a model in which individuals arrive to a contaminated area according to a Yule process with immigration.</p>		
<p>Reproduced by NATIONAL TECHNICAL INFORMATION SERVICE U S Department of Commerce Springfield VA 22151</p>		

DD FORM 1 NOV 6 1473 (PAGE 1)

S/N 0101-807-6801

UNCLASSIFIED

Security Classification

UNCLASSIFIED
Security Classification

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
YULE PROCESS						
ORDER STATISTICS						
EXPONENTIAL DISTRIBUTION						
OPTIMAL EVACUATION						

DD FORM 1 NOV 68 1473 (BACK)
(PAGE 2)

Security Classification

i-a

EVACUATION OF A YULE PROCESS WITH IMMIGRATION

MARK BROWN, SHELDON ROSS, RICHARD SHORROCK

TECHNICAL REPORT NO. 19

APRIL 29, 1974

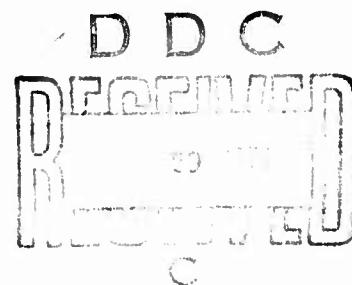
PREPARED UNDER CONTRACT N00014-67-A-0214-0015

(NR-042-267)

Herbert Solomon, Project Director

**Reproduction in Whole or in Part is Permitted for
any Purpose of the United States Government**

**DEPARTMENT OF STATISTICS
THE GEORGE WASHINGTON UNIVERSITY
WASHINGTON, D.C. 20006**



EVACUATION OF A YULE PROCESS WITH IMMIGRATION

by

Mark Brown, Sheldon Ross, Richard Shorrock

1. Introduction

In the second section of this paper we obtain a correspondence between the joint distribution of the epochs of birth in a Yule process and the order statistics of a sequence of independent and identically distributed (i.i.d.) exponential random variables. This correspondence is then used to obtain simple proofs of some Yule process results. In particular a simple proof of a result due to Neuts and Resnick [1] concerning the conditional distribution of the birth epochs given the number of such epochs in $(0, t)$ is presented.

In the final three sections we deal with the problem that motivated this paper. Namely if a population of individuals arrives at a contaminated geographical area in accord with a Yule process with immigration, and if all people in the area are to be evacuated at time T because of the unhealthy conditions in the area, then if we are allowed an additional evacuation time in $(0, T)$ how can we choose it in an optimal manner?

2. The Yule Process

The Yule process is a pure birth process with birth parameters $\lambda_j = j\lambda$, $j \geq 0$. In other words, each individual in the population independently gives birth in accordance with an exponential distribution with rate λ . We assume that there is a single person in the population at time $t=0$. Define T_i , $i \geq 1$, to be the time of the i -th birth, and let $N(t)$ equal the number of births in $(0, t)$.

Let x_1, \dots, x_k be k i.i.d. exponential random variables with rate λ and let $x_{(i)}$ equal the i -th smallest of x_1, \dots, x_k for $i=1, \dots, k$. It directly follows, from the lack of memory of exponential random variables, that

$$(C) \quad T_1, T_2, \dots, T_k$$

has the same joint distribution as

$$(C) \quad x_{(k)} - x_{(k-1)}, x_{(k)} - x_{(k-2)}, \dots, x_{(k)}.$$

We shall refer to this fact as the correspondence C.

As an application of the correspondence we note that

$$P\{T_k < t\} = P\{x_{(k)} < t\} = [1 - e^{-\lambda t}]^k$$

and thus

$$\begin{aligned} P\{N(t) = k\} &= P\{N(t) \geq k\} - P\{N(t) \geq k + 1\} \\ &= P\{T_k < t\} - P\{T_{k+1} < t\} \\ &= e^{-\lambda t} [1 - e^{-\lambda t}]^k \end{aligned}$$

which, of course, is just the well known result that the transition probabilities in the Yule process are given by

$$P_{1,k+1}(t) = e^{-\lambda t} [1 - e^{-\lambda t}]^k$$

It easily follows from the above that

$$\lim_{t \rightarrow \infty} P\{e^{-\lambda t} N(t) > x\} = e^{-x}$$

and, in fact, it can be shown, by regarding T_k as a weighted sum of i.i.d. exponential random variables, that $T_k / \log k \rightarrow \frac{1}{\lambda}$ with probability 1, and this can easily be shown to yield that

$$\frac{\log N(t)}{t} \rightarrow \lambda \text{ with probability 1.}$$

The correspondence C immediately yields the following

Proposition 1: Given $T_k = t, T_1, \dots, T_{k-1}$ have the same joint distribution as the order statistics from a sample

of $k-1$ i.i.d. random variables having density function

$$f(x) = \begin{cases} \frac{\lambda e^{-\lambda(t-x)}}{1-e^{-\lambda t}} & 0 \leq x \leq t \\ 0 & \text{otherwise} \end{cases}$$

Proof: From the correspondence C we note that the conditional distribution of T_1, \dots, T_{k-1} given $T_k=t$ is the same as the conditional distribution of $t-X_{(k-1)}, \dots, t-X_{(1)}$ given $t-X_{(k)}$. As $f(x)$ is just the conditional density of $t-Y$ given that $Y < t$, where Y is an exponential random variable with rate λ , the result follows.

While Proposition 1 yields the conditional distribution of times of birth given $T_k=t$, it is more useful to obtain their conditional distribution given information about $N(t)$. To obtain this, note that the conditional distribution of T_{k+1} given $N(t)=k, T_1, \dots, T_k$ is the same as the conditional distribution of T_{k+1} given $N(t)=k$. Therefore, given $N(t)=k$, T_{k+1} is conditionally independent of T_1, \dots, T_k , and thus the conditional distribution of T_1, \dots, T_k given $N(t)=k$ and $T_{k+1}=t$ is the same as the conditional distribution of T_1, \dots, T_k given $N(t)=k$. Hence, as the event that $N(t)=k$ and $T_{k+1}=t$ is, with probability 1, equivalent to the event that $T_{k+1}=t$, we obtain from Proposition 1.

Proposition 2: Given $N(t)=k, T_1, \dots, T_k$ are distributed as the order statistics from a sample of size k from a population having density

$$f(x) = \begin{cases} \frac{\lambda e^{-\lambda(t-x)}}{1-e^{-\lambda t}} & 0 \leq x \leq -1 \\ 0 & \text{otherwise} \end{cases}$$

Proposition 2 can be used in establishing results about the Yule process in the same way that the corresponding result for Poisson Processes is used. For example, consider the random variable $Y(t)$ defined by

$$Y(t) = \sum_{i=1}^{N(t)} g_i(t-T_i)$$

where $g_i(\cdot)$, $i \geq 1$, are i.i.d. random functions that are independent of the Yule process. (Yule shot noise would probably be an appropriate name for the stochastic process $\{Y(t), t \geq 0\}$). Letting $\psi_t(u) = E[e^{iuY(t)}]$ be the characteristic function of $Y(t)$, we obtain $\psi_t(u)$ by conditioning on $N(t)$ as follows.

$$E[e^{iuY(t)}] = E\{E[e^{iuY(t)} | N(t) = k]\}$$

Now, since given $N(t)=k$, the unordered set $\{T_1, \dots, T_k\}$ are distributed as i.i.d. random variables with density

given in Proposition 2 it follows that

$$\begin{aligned} E\left[e^{iuY(t)} | N(t)=k\right] &= \left[\frac{\lambda}{1-e^{-\lambda t}} \int_0^t e^{-\lambda(t-x)} \psi_{g(t-x)}(u) dx \right]^k \\ &= \left[\frac{\lambda}{1-e^{-\lambda t}} \int_0^t e^{-\lambda x} \psi_{g(x)}(u) dx \right]^k \end{aligned}$$

where

$$\psi_{g(x)}(u) = E\left[e^{iug(x)}\right].$$

Hence,

$$\begin{aligned} \psi_t(u) &= e^{-\lambda t} \sum_{k=0}^{\infty} \left[\int_0^t \lambda e^{-\lambda x} \psi_{g(x)}(u) dx \right]^k \\ &= \frac{e^{-\lambda t}}{1 - \int_0^t \lambda e^{-\lambda x} \psi_{g(x)}(u) dx} \end{aligned}$$

Differentiation yields, assuming the moments exist, that

$$E[Y(t)] = \lambda e^{\lambda t} \int_0^t e^{-\lambda x} E[g(x)] dx$$

$$\text{Var}[Y(t)] = \lambda e^{\lambda t} \int_0^t e^{-\lambda x} E[g^2(x)] dx + \lambda^2 e^{2\lambda t} \left[\int_0^t e^{-\lambda x} E[g(x)] dx \right]^2$$

3. Optimal Evacuation of a Yule Process with Immigration

Individuals arrive at a geographical area that is initially empty, in accordance with a pure birth process with birth parameters $\lambda_j = j\lambda + \theta$, $j \geq 0$. Due to contamination, this geographical area is unsafe for its population and at some fixed time T in the future everyone in the area will be evacuated and no further immigration will be allowed.

If an individual spends a total time of x units in the area before being evacuated then let us suppose that a cost of $g(x)$ is incurred (by society). We are interested in deriving an expression for the total expected cost incurred by time T . To do so let us at first assume that a single individual is present at time $t=0$ and that the immigration parameter θ equals 0. In this case, we have by (2) that the expected cost incurred by time T is

$$C_1(T) = g(T) + \lambda e^{\lambda T} \int_0^T e^{-\lambda x} g(x) dx$$

When there is a single individual present at time 0 but $\theta \neq 0$ then, since each immigrant can be thought of as starting his own Yule process when he arrives, we obtain by conditioning on the total number of immigrants that arrive in $(0, T)$ that $C_2(T)$, the expected cost incurred, is given by

$$C_2(T) = C_1(T) + \theta T \int_0^T C_1(T - x) \frac{dx}{T}$$

$$= C_1(T) + \theta \int_0^T C_1(T-x) dx$$

Finally, when there is no one present at time $t=0$, we obtain, by conditioning on the arrival time of the initial immigrant, that the total expected cost incurred is

$$C(T) = \int_0^T C_2(T-s) \theta e^{-\theta s} ds$$

In the special case in which the loss incurred is equal to the time spent in the contaminated area, i.e., in the case $g(x) = x$, a simple computation yields that

$$C(T) = \frac{\theta}{\lambda^2} [e^{\lambda T} - \lambda T - 1] \quad \text{when } g(x) = x$$

3. Optimal Evacuation of a Yule Process with Immigration

Individuals arrive at a geographical area that is initially empty, in accordance with a pure birth process with birth parameters $\lambda_j = j\lambda + \theta$, $j \geq 0$. Due to contamination, this geographical area is unsafe for its population and at some fixed time T in the future everyone in the area will be evacuated and no further immigration will be allowed..

If an individual spends a total time of x units in the area before being evacuated then let us suppose that a cost of $g(x)$ is incurred (by society). We are interested in deriving an expression for the total expected cost incurred by time T . To do so let us at first assume that a single individual is present at time $t=0$ and that the immigration parameter θ equals 0. In this case, we have by

(2) that the expected cost incurred by time T is

$$C_1(T) \equiv g(T) + \lambda e^{\lambda T} \int_0^T e^{-\lambda x} g(x) dx$$

When there is a single individual present at time 0 but $\theta \neq 0$ then, since each immigrant can be thought of as starting his own Yule process when he arrives, we obtain by conditioning on the total number of immigrants that arrive in $(0, T)$ that $C_2(T)$, the expected cost incurred, is given by

$$C_2(T) = C_1(T) + \theta T \int_0^T C_1(T-x) \frac{dx}{T}$$

$$= C_1(T) + \theta \int_0^T C_1(T-x) dx$$

Finally, when there is no one present at time $t=0$, we obtain, by conditioning on the arrival time of the initial immigrant, that the total expected cost incurred is

$$C(T) = \int_0^T C_2(T-s) \theta e^{-\theta s} ds$$

In the special case in which the loss incurred is equal to the time spent in the contaminated area, i.e., in the case $g(x) = x$, a simple computation yields that

$$C(T) = \frac{\theta}{\lambda^2} [e^{\lambda T} - \lambda T - 1] \quad \text{when } g(x) = x$$

4. Optimal Intermediate Evacuation Time

Suppose now that an intermediate evacuation time τ , $0 \leq \tau \leq T$, at which time everyone present in the area would be evacuated, is to be chosen. The area would then again fill up with individuals between times τ and T , and, at T , the final evacuation would be made and the area would be permanently sealed off. The problem is to choose τ so as to minimize the total expected cost incurred by time T .

Let us say that we are in state (t, x_1, \dots, x_n) if the intermediate evacuation has not yet been made and if there are n individuals of respective ages x_1, \dots, x_n present at time $T-t$. If we are in state (t, x_1, \dots, x_n) and we decide to evacuate immediately then our total expected cost is

$$C(t) + \sum_{i=1}^n g(x_i) ;$$

on the other hand if we wait a fixed additional time ϵ before evacuating then our total expected cost is

$$C(t - \epsilon) + \sum_{i=1}^n g(x_i + \epsilon) + (n\lambda + \theta) \int_0^\epsilon g(s) ds + o(\epsilon)$$

Hence evacuating in state (t, x_1, \dots, x_n) is better than waiting a fixed additional time ϵ before evacuating if

$$C(t) + \sum_{i=1}^n g(x_i) \leq C(t - \epsilon) + \sum_{i=1}^n g(x_i + \epsilon) + (n\lambda + \theta) \int_0^\epsilon g(s) ds + o(\epsilon)$$

or, equivalently, if

$$\frac{C(t) - C(t - \epsilon)}{\epsilon} \leq \sum_{i=1}^n \frac{g(x_i + \epsilon) - g(x_i)}{\epsilon} + (n\lambda + \theta) \int_0^\epsilon \frac{g(s)ds}{\epsilon} + \frac{o(\epsilon)}{\epsilon}$$

Letting $\epsilon \rightarrow 0$ shows that if $\int_0^\epsilon g(s)ds = o(\epsilon)$, then evacuating when in state (t, x_1, \dots, x_n) is better than waiting an infinitesimal additional amount of time if

$$C'(t) \leq \sum_{i=1}^n g'(x_i)$$

We shall assume that $g(s)$ is such that $\int_0^\epsilon g(s)ds = o(\epsilon)$, and thus the above defines the infinitesimal look ahead rule (see [2]) - it says to evacuate when n individuals of respective ages, x_1, \dots, x_n are present at time $T-t$ if

$$\sum_{i=1}^n g'(x_i) \geq C'(t)$$

Proposition 3:

Assume (i) $g(0) = 0$

(ii) $g'(x) \geq 0$

(iii) $g''(x) \geq 0$

(iv) $\lim_{\epsilon \rightarrow 0} \int_0^\epsilon \frac{g(s)ds}{\epsilon} = 0$

Then the infinitesimal look ahead rule is optimal. That is, assuming (i), (ii), (iii), (iv), the intermediate evacuation time that minimizes the total expected cost

is the one that evacuates at the smallest t such that

$$\sum_{i=1}^{N(t)} g'(x_i) \geq C'(T - t)$$

where $N(t)$ is the number of individuals in the area at time t and their respective ages are $x_1, \dots, x_{N(t)}$. In other words, the intermediate evacuation should be made the first time that the instantaneous rate of cost increase due to those present in the area is greater than the instantaneous cost increase if an evacuation is made.

Proof: It was proven in [2] that if the set of states that the infinitesimal look ahead rule tells us to evacuate at is a closed set of states, in the sense that once we are in one of these states then we can never leave the set, then this rule is optimal. In our case since $\sum_{i=1}^{N(t)} g'(x_i)$ cannot decrease as t increases (until an evacuation occurs) we will establish the result if we can show that $C'(T - t)$ is a nonincreasing function of t . That is we must show that $C(T)$ is a convex function of T .

Now

$$C'_1(T) = g'(T) + \lambda g(T) + \lambda^2 e^{\lambda T} \int_0^T e^{-\lambda x} g(x) dx \geq 0$$

and thus

$$C''_1(T) = g''(T) + \lambda g'(T) + \lambda^3 e^{\lambda T} \int_0^T e^{-\lambda x} g(x) dx + \lambda^2 g(T) \geq 0$$

Also,

$$c_2(T) = c_1(T) + \theta \int_0^T c_1(y) dy$$

and thus

$$c_2'(T) = c_1'(T) + \theta c_1(T)$$

Differentiating again yields

$$c_2''(T) = c_1''(T) + \theta c_1'(T) \geq 0$$

Finally,

$$c(T) = \theta e^{-\theta T} \int_0^T e^{\theta y} c_2(y) dy$$

$$c'(T) = -\theta^2 e^{-\theta T} \int_0^T e^{\theta y} c_2(y) dy + \theta c_2(T)$$

$$c^{11}(T) = -\theta^2 c_2(T) + \theta^3 e^{-\theta T} \int_0^T e^{\theta y} c_2(y) dy + \theta c_2'(T)$$

Hence, we need to show that

$$\theta^2 e^{-\theta T} \int_0^T e^{\theta y} c_2(y) dy + c_2'(T) \geq \theta c_2(T)$$

Integrating by parts yields that

$$\begin{aligned} \int_0^T \theta e^{\theta y} c_2(y) dy &= c_2(y) e^{\theta y} \Big|_0^T - \int_0^T e^{\theta y} c_2'(y) dy \\ &= c_2(T) e^{\theta T} - \int_0^T e^{\theta y} c_2'(y) dy \end{aligned}$$

and thus we need to show that

$$\theta C_2(T) - \theta e^{-\theta T} \int_0^T e^{\theta y} C_2'(y) dy + C_2'(T) \geq \theta C_2(T)$$

or, equivalently, that

$$C_2'(T) \geq \theta e^{-\theta T} \int_0^T e^{\theta y} C_2'(y) dy$$

Now, since $C_2''(y) \geq 0$, it follows that

$$\begin{aligned} \theta e^{-\theta T} \int_0^T e^{\theta y} C_2'(y) dy &\leq \theta e^{-\theta T} C_2'(T) \int_0^T e^{\theta y} dy \\ &= C_2'(T) (1 - e^{-\theta T}) \\ &\leq C_2'(T) \end{aligned}$$

which proves the result.

Hence, in the special case $g(x) = x$ where we are interested in minimizing the total expected time spent in the area by all individuals, it follows that it is optimal to evacuate at time t if

$$N(t) \geq \frac{\theta}{\lambda} [e^{\lambda(T-t)} - 1]$$

When $\lambda = 0$, this reduces to evacuating for the first t such that $N(t) \geq \theta(T-t)$, a result first established in [3].

The optimal evacuation time given by Proposition 3 was derived under the assumption that we are, at all times, aware of the number of individuals in the area. However, it

is quite possible that we might not have any information about the population size before the evacuation is made. In this case we would thus be forced to choose a constant time for the intermediate evacuation. Since evacuating at the fixed time t leads to an expected cost of $C(t) + C(T - t)$ it follows from convexity that, under the assumptions of Proposition 3, the optimal constant evacuation time is $t = T/2$. In fact this result easily generalizes to the case where there are n constant intermediate evacuations to be made. Under the assumptions of Proposition 3, it again follows by convexity that the optimal times would be $T/n+1, 2T/n+1, \dots, nT/n+1$.

5. The Unknown Parameter Case

Up to this point we've assumed that λ and θ were known parameters. However, it may well be the case that one or even both of these values are unknown. What procedure should we employ in this case?

To make explicit its dependence on λ and θ , let us write $C(T, \lambda, \theta)$ for $C(T)$. From Proposition 3, it would seem that a reasonable procedure would be to evacuate at the first time t such that

$$\sum_{i=1}^{N(t)} g'(x_i) \geq \frac{2}{2T} C(T - t, \hat{\lambda}(t), \hat{\theta}(t))$$

where $x_1, \dots, x_{N(t)}$ are the respective ages of those in the area at time t , and $\hat{\lambda}(t)$ and $\hat{\theta}(t)$ are respective estimates of λ and θ at time t . (Naturally we are assuming that

$g(x)$ satisfies the conditions of Proposition 3). Thus we are led to determining estimates for λ and θ .

If we let $I(t)$ denote the number of immigrants to the area in $[0, t]$ then $\hat{\theta} = \frac{I(t)}{t}$ is the obvious estimate for θ .

Letting $A(t)$ denote the sum of the amount of time that has been spent in the area by those present at time t , and letting $B(t)$ denote the number of births that occur in the area in $[0, t]$, it turns out that the maximum likelihood estimate of λ is $\hat{\lambda}(t) = \frac{B(t)}{A(t)}$. (This is the same as the total time on test statistic used in exponential sampling schemes). Since the increase from birth to birth of the sum of the times spent in the area are independent and identically distributed exponential random variables with mean $1/\lambda$, it easily follows that, when $\theta > 0$, $\hat{\lambda}(t) \rightarrow \lambda$ as $t \rightarrow \infty$ with probability 1.

Thus, for instance, a reasonable procedure for determining the intermediate evacuation time in the case where $g(x) = x$, would be to evacuate the first time t for which

$$N(t) \geq \frac{I(t) A(t)}{t B(t)} \quad \exp \left[\left\{ \frac{B(t)}{A(t)} (T - t) \right\} - 1 \right]$$

It should be noted that the intermediate evacuation time defined above may not be optimal in any sense. However it is clearly the procedure suggested by Proposition 3. It is interesting to note that when $B(t) = 0$ the above reduces to evacuating at

the first t for which

$$N(t) \geq \frac{I(t)(T-t)}{t}$$

But since $N(t) = I(t) + B(t)$, the above inequality is equivalent to

$$t \geq T/2$$

Hence if no births occur by time $T/2$ then the intermediate evacuation should be made at $T/2$.

References

- [1] Neuf, M. and Resnick, S., "On the Times of Births in a Linear Birthprocess," Journal of the Australian Math Society, Vol. XII, Part 4, 473-475.
- [2] Ross, S., "Infinitesimal Look Ahead Stopping Rules," Annals of Mathematical Statistics, 1971, Vol. 42, No. 1, 297-303.
- [3] Ross, S., "Optimal Dispatching of a Poisson Process," Journal of Applied Probability.